

DOI: 10.19650/j.cnki.cjsi.J2007207

基于改进 DQN 网络的滚动轴承故障诊断方法*

康守强, 刘 哲, 王玉静, 王庆岩, 兰朝凤

(哈尔滨理工大学电气与电子工程学院 哈尔滨 150080)

摘 要:针对实际中滚动轴承正常和故障状态下的振动数据不平衡,且故障诊断准确率不高的问题,基于深度强化学习,提出一种改进深度 Q 网络(DQN)的滚动轴承故障诊断方法。该方法将振动信号进行短时傅里叶变换,构建时频图样本集;提出把 K-means 算法中样本到中心点的距离作为回报值的偏置,以不平衡比为基准,为训练集构建具有个性化的回报函数,同时引入残差网络(Resnet-18)实现特征的深层提取;智能体将新的回报函数和时频图作为输入,在每个时间步长执行诊断动作,判断并返回回报值;最终,智能体学会不平衡数据下的故障诊断策略。实验表明,所提改进的诊断模型相比本文对比的其他方法在不平衡下提高了 5%~8%;同时不平衡且变负载情况下也表现突出,不平衡指标得分达到了 0.982 左右,具有较好的泛化性。

关键词:滚动轴承;不平衡;K-means;故障诊断;深度强化学习

中图分类号: TN911.7 TH165.3 **文献标识码:** A **国家标准学科分类代码:** 460.4099

A fault diagnosis method of rolling bearing based on the improved DQN network

Kang Shouqiang, Liu Zhe, Wang Yujing, Wang Qingyan, Lan Chaofeng

(School of Electrical and Electronic Engineering, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: Under normal and fault states in practice, rolling bearing vibration data are imbalanced and the fault diagnosis accuracy is low. Based on the deep reinforcement learning, an improved deep Q network (DQN) fault diagnosis method for rolling bearing is proposed. The short time Fourier transform is performed on the vibration data to establish sample sets of time-frequency graph. The distance between the sample and the center point in the K-means algorithm is used as the bias of the return value. The imbalance ratio is utilized as the benchmark to formulate a personalized reward function for the training set. Meanwhile, the residual network (Resnet-18) is used to realize the deep extraction of features. In which, the agent takes the new reward function and time-frequency graph as input. The diagnosis action is executed at each time step. And the reward is judged and returned. Finally, the agent learns the fault diagnosis strategy under imbalanced data. Compared with other methods, experimental results show that the improved diagnostic model is improved by 5% to 8% under imbalanced conditions. At the same time, it also performs outstandingly under imbalanced and variable load conditions. The imbalanced index score can reach about 0.982, which shows better generalization.

Keywords: rolling bearing; imbalanced; K-means; fault diagnosis; deep reinforcement learning

0 引 言

滚动轴承作为旋转机械设备的重要部件之一,被广泛应用于工业领域^[1,2],对其进行故障诊断有利于预防设备事故发生^[3]。滚动轴承在实际工作中,大部分时间处在正常运行状态,采集到的正常状态的振动信号样本远

比故障状态的要多,训练得到的模型也对多数类样本敏感,使少数类样本难以被识别。近年来,滚动轴承振动数据不平衡的故障诊断研究受到学者的广泛关注,对处理实际的工业数据具有很强的应用价值和现实意义。

在数据分布不平衡情况下,主要有两类解决方法^[4]:从数据角度,通过对训练样本进行重采样,使训练样本的分布达到平衡;从算法角度,不改变训练数据集分布,通

收稿日期:2020-12-08 Received Date: 2020-12-08

* 基金项目:国家自然科学基金(51805120)、黑龙江省自然科学基金(LH2019E058)、黑龙江省普通高校基本科研业务专项资金(LGYC2018JC022)资助

过调整分类算法使少数类样本得到更多的关注。

数据角度的方法包括过采样、欠采样和混合采样。过采样方法通过增加少数类样本使训练集达到平衡,文献[5]提出 K^* -信息量近邻域过采样方法,有效地解决了轴承故障样本失衡问题,取得了不错的效果;欠采样方法通过舍弃部分多数类数据实现平衡训练集,文献[6]为了解决大数据获取和诊断效率之间的平衡问题,提出在数据采集阶段用欠采样方法减少大数据,且在滚动轴承故障诊断中验证了其有效性;混合采样是二者的结合,文献[7]引入主曲线和造粒分布模拟数据的总体分布特征,以进行可靠的过采样和欠采样,提出了基于极限学习机在线顺序预测方法,在不平衡故障诊断任务中的准确率高达 95%~97%。这些方法通过改变数据的不平衡分布,从而达到提升分类性能的目的。

算法角度的方法主要有分类阈值调整法、基于 Boosting 的集成学习和基于代价的敏感学习等。对于分类阈值调整的方法,通过调整分类器的分类边界的阈值,改变对不同输出概率的类别判定。文献[8]提出了一种加权 softmax 损失来解决不平衡分类问题,并且在 3 个具有不同不平衡度的轴承数据集进行验证,可以有效地处理分类不平衡问题。基于集成学习的思想,文献[9]在保证少数类样本充分参与训练的前提下,训练多个基分类器,提出多分类器集成加权均衡分布适配的滚动轴承寿命阶段识别方法, F-score 均分达到了 0.73,有效地识别了少数类样本。代价敏感学习为少数类错分实例赋予更大的错分代价,是一种常用技巧。文献[10]在设计加权损失函数优化不平衡数据分布的同时,也结合随机欠采样平衡训练样本,在 PHM2015 工厂故障事件数据集中进行了验证,准确率比其他基准方法高出 2%~3%。

数据不平衡现象广泛存在于工程实际中,但针对滚动轴承振动数据不平衡的故障诊断文献较少,是当下研究的一个热点。同时,上述两类解决办法都具有一定的局限性,数据角度的改进方法虽然平衡了数据集,但会改变原始数据分布,不能准确把握数据的真实情况;算法角度的改进方法是让模型等价地学习不平衡样本,也属于对特征样本的学习。而数据不平衡问题是客观的,当少数类样本的总体特征空间无法估计时,数据和算法角度的改进方法都有一定的不足。

针对上述问题,文献[11]另辟蹊径,将分类问题模拟为智能体(agent)连续决策的过程,运用深度强化学习中的探索-利用机制,建立了基于深度 Q 网络(deep Q-network, DQN)的不平衡分类模型。深度强化学习作为一种成熟框架,由于其独有的反馈机制,在分类问题上也有着较为广泛的应用。文献[12]提出基于深度强化学习的特征选择方法,将特征选择和分类定义为连续决策问题,在智能体每次决策时,通过已选取的特征决定是

否再次请求一个特征,在 mnist 等多个公开数据集中验证了该方法的有效性。文献[13]针对数据中含有噪声的问题,提出利用深度强化学习筛选高质量的句子样例,实现了噪声文本的样本筛选与关系分类。文献[14]首次将分类任务模拟成强化学习中的顺序决策过程,提出了一种基于强化学习的分类任务解决方案,并在 UCI 八种医学疾病数据集中的准确率达到 87.4%。文献[15]针对原有深度神经网络需要人工调参和专家经验的问题,提出利用深度强化学习算法 DQN 建立端到端的诊断模型,并在滚动轴承和液压泵的数据集上进行了验证,准确率在 90%~94%之间。

上述部分文献利用强化学习做特征选择,本质上以深度学习为主、强化学习为辅,在面对不平衡数据集时略显乏力;而文献[11]直接将分类任务模拟为深度强化学习中的连续决策,通过赋予不同的回报值,使模型适应不平衡数据集。但仅根据不平衡比来设定回报值,只考虑了类间的不平衡,对于类内的数据无法区分重要度。同时,上述文献将深度强化学习应用在旋转机械的故障诊断时,没有讨论变负载、不平衡问题且在特征提取和回报函数上仍需深入研究,有一定的提升空间。

因此,为提高滚动轴承数据不平衡的故障诊断准确率,本文提出一种改进 DQN 的滚动轴承故障诊断方法。该方法把故障诊断过程模拟成 DQN 模型中的顺序决策过程,模型将振动信号构建的二维图像数据集和根据 K-means 算法设计的回报函数作为输入,通过 Resnet-18 提取深层特征,在每个时间步长执行诊断动作,判断并返回回报值,实现滚动轴承振动数据不平衡的多状态智能故障识别。

1 理论背景与模型改进

1.1 DQN 理论

DQN 的核心理论就是通过一个智能体(agent)来感知环境状态,并且在与环境交互的过程中得到最大收益的目标。

在智能体的判断过程中,策略函数 π 收到一个状态样本 s_t ,并依据一定的概率 P 返回该状态的动作 a_t (故障类型)。

$$\pi(a|s) = P(a_t = a | s_t = s) \quad (1)$$

智能体的目标就是尽可能正确地识别训练集样本,智能体在正确识别样本时获得正奖励 R ,通过最大化累积奖励来实现其目标 G_t :

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (2)$$

式中: γ 为加权值。

在 t 时刻所做的决定在 $t+1$ 时刻才有回报,其中, γ

决定了未来回报的重要程度,当 $\gamma = 0$ 时,agent 总是着眼立即回报;当 $\gamma = 1$ 时,agent 总是着眼于长远回报。那么状态 s 下的期望回报 V 为:

$$V(s) = E(G_t | s_t = s) = E[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots] = E[R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + \dots)] = E[R_{t+1} + \gamma V(s_{t+1})] \quad (3)$$

式(3)又称 Bellman 方程,展现了当前回报和未来回报之间的递归关系。同时,引入 Q 函数,返回该动作在该状态下的预期未来奖励:

$$Q^\pi(s_t, a_t) = E[r_{t+1} + \gamma Q^\pi(s_{t+1}, a_{t+1}) | s_t, a_t] \quad (4)$$

$$V(s) = \sum_{a \in A} \pi(a | s) Q(s, a) \quad (5)$$

式中:策略记号 π 是策略函数;动作 a 属于动作集合 A 。

在策略 π 下, $V(s)$ 是所有可能的动作的长期回报的期望,最优的策略对应最优的 V 和 Q 值,即:

$$V^*(s) = \max_{\pi} V_{\pi}(s) = \max_a Q^*(s, a) \quad (6)$$

$$Q^*(s, a) = \max_{\pi} Q_{\pi}(s, a) = E[r_t + \gamma \max_a Q^*(s_{t+1}, a_{t+1}) | s_t, a_t] \quad (7)$$

Q 函数利用一个表格来存储 (s, a, r, s') 队列。当空间状态维度较大时,将 Q 学习算法和深度学习结合起来,形成了深度 Q 网络^[16]。DQN 模型为 Q 学习算法添加“深度”,在准确率方面有了质的提高,广泛应用在许多领域。DQN 的训练流程如图 1 所示。

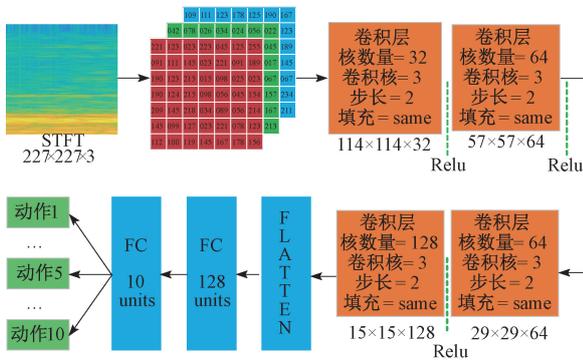


图 1 深度 Q 网络流程

Fig. 1 Flow of the deep Q -network

在 DQN 模型中用深度神经网络来逼近动作值函数 Q^* , 然后对网络进行训练,更新网络参数 θ 以最小化损失函数,损失函数如式(8)所示。

$$L(\theta_i) = E_{s,a,r,s'} [(y_i - Q(s, a | \theta_i))^2] \quad (8)$$

其中, y_i 定义为:

$$y_i = r + \gamma Q(s', a' | \theta_i^-) \quad (9)$$

式中: s' 是 s 的下一个状态; a' 是智能体在状态 s' 下执行的动作。

损失函数 L 对参数 θ 求导为:

$$\nabla_{\theta_i} L(\theta_i) = E_{s,a,r,s'} [(r + \gamma Q(s', a' | \theta_i^-) - Q(s, a | \theta_i)) \nabla_{\theta_i} Q(s, a | \theta_i)] \quad (10)$$

通过式(10)最小化损失函数求得最优 Q^* 函数的同时,式(8)获得最大累积奖励,得到最优的分类策略 $\pi: S \rightarrow A$ 。

1.2 改进 DQN 网络的滚动轴承故障诊断方法

1) 修改网络提升深度

为了增加网络的深度,减少模型的过拟合问题,提出引入 Resnet-18 网络^[17] 搭建 DQN 网络模型。同时为了使结构在计算上更简便,不通过全连接层输出全部 Q 值,而是输出对应状态和动作的 Q 值,加快了网络训练速度。修改后网络的基本流程如图 2 所示,具体的改动设置如 3.2 节表 2 所示。

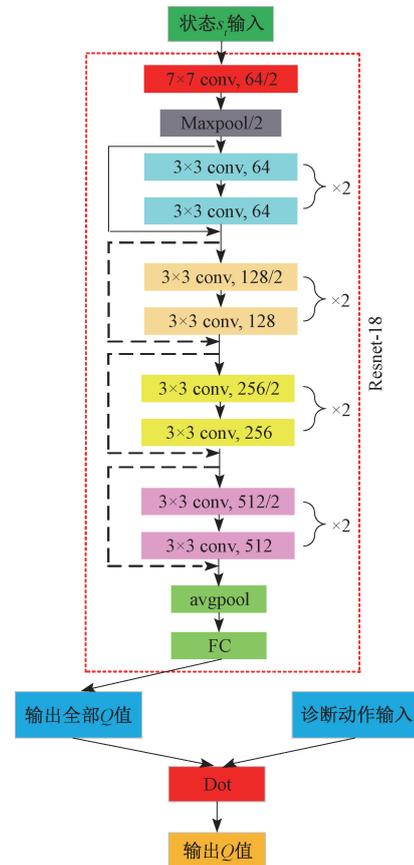


图 2 修改后网络的基本流程

Fig. 2 The basic flow of the modified network

2) 权衡样本的距离量化奖励函数

(1) 根据训练集的不平衡比设计奖励函数

少数类样本难以在不平衡数据集中正确识别,为了更好地识别少数类的样本,模型应给予其更多的关注,当智能体遇到少数类样本时,获得较大的奖励或惩罚。奖励函数定义如式(11)所示。

$$R(s_i, a_i, y_i) = \begin{cases} +1, & a_i = y_i \& s_i \in D_N \\ -1, & a_i \neq y_i \& s_i \in D_N \\ \rho, & a_i = y_i \& s_i \in D_F \\ -\rho, & a_i \neq y_i \& s_i \in D_F \end{cases} \quad (11)$$

式中: $1/\rho \in [0, 1]$, $\rho = D_N/(D_F/9)$ 表示不平衡比例; D_N 是多数类正常样本; D_F 是少数类故障样本; y_i 是状态 s_i 下的样本标签。当智能体正确/错误的对多数类样本进行分类时, 奖励值为 $1/-1$; 当智能体正确/错误的对少数类样本进行分类时, 奖励值为 $\rho/-\rho$ 。

(2) 通过权衡样本的欧式距离量化奖励函数

为了更好的识别少数类样本, 本文利用 K 均值聚类对奖励函数进行量化, 使得每个样本都有其各自的奖励值。K 均值聚类算法是一种无监督的聚类算法, 以距离作为数据对象间相似性度量的标准, 相似度与间距大小成反比, 最终将数据分为 k 类簇。

将 K 均值聚类的中心点作为不同类别回报值的基本点, 其中心点的回报值如式 (11) 所示; 类别之内利用每个样本与所属簇中心点之间的欧氏距离来对回报函数进行量化, 具体步骤:

训练集每张图片的维度较大, 首先采用主成分分析法 (principal components analysis, PCA) [18] 来对数据进行降维, 然后将数据集聚类成 10 个簇 $C = C_1, C_2, \dots, C_{10}$, 随机选取 k 个簇中心, 即 $\{\mu_1, \mu_2, \dots, \mu_k\}$ 。更新样本所对应的簇, 最小化损失函数为:

$$E = \sum_{i=1}^k \sum_{s_j \in C_i} \|x - \mu_i\|^2 \quad (12)$$

式中: s_i 为训练集样本; μ_i 为簇 C_i 的中心点; k 为簇的类别数。

更新各个簇中心 $\mu_1, \mu_2, \dots, \mu_k$ 为:

$$\mu_i = \frac{1}{C_i} \sum_{s_j \in C_i} s_j \quad (13)$$

遍历所有的簇分类找到以上问题的最优解, 直到簇标签达到收敛精度为止, 否则继续重复最小化损失函数和更新各个簇中心两个步骤。最终确定 k 个簇中心点, 即 k 个分类所对应的奖励函数值中心点, 将每个样本与中心点的距离进行归一化处理。

$$Dis(s_i) = \frac{\|s_i - \mu_i\|^2}{\|s_i - \mu_i\|_{\max}^2 - \|s_i - \mu_i\|_{\min}^2} \quad (14)$$

通过比较不同类别所对应的奖励函数值与同类别中每个样本与中心的距离, 对奖励函数值进行量化。

$$z(s_i) = R(s_i, a_i, y_i, \mu_i) = \begin{cases} 1 - Dis(s_i), & s_i \in D_N \& y_i = C_i \\ \rho - Dis(s_i), & s_i \in D_F \& y_i = C_i \end{cases} \quad (15)$$

3) 奖励函数影响分析

假设多样本和少样本分别表示为 s^+ 和 s^- , 他们的目

标 Q 值分别表示为 y^+ 和 y^- , 通过式 (9) 和 (15), 目标 Q 值可以表示为:

$$y^+ = (-1)^{1-B(a=y)} z(s, a) + \gamma \max_{a'} Q(s', a') \quad (16)$$

$$y^- = \rho(-1)^{1-B(a=y)} z(s, a) + \gamma \max_{a'} Q(s', a') \quad (17)$$

式中: $B(x)$ 是判定函数, 返回动作和标签的判定结果。

DQN 的损失函数重写为多样本损失函数 $L_+(\theta_i)$ 和少样本损失函数 $L_-(\theta_i)$, 其导数分别为:

$$\nabla_{\theta_i} L_+(\theta_i) = E_{s, a, r, s'}^N [(y_i^+ - Q(s^+, a | \theta_i)) \nabla_{\theta_i} Q(s^+, a | \theta_i)] \quad (18)$$

$$\nabla_{\theta_i} L_-(\theta_i) = E_{s, a, r, s'}^F [(y_i^- - Q(s^-, a | \theta_i)) \nabla_{\theta_i} Q(s^-, a | \theta_i)] \quad (19)$$

将式 (16)、(17) 带入 $L_+(\theta_i)$ 和 $L_-(\theta_i)$ 得:

$$\begin{aligned} \nabla_{\theta_i} L(\theta_i) &= E_{s, a, r, s'}^{N+F} [(y^{+/-} - Q(s, a | \theta_i)) \times \\ &\nabla_{\theta_i} Q(s, a | \theta_i)] = E_{s, a, r, s'}^{N+F} \gamma \max_{a'} Q(s', a' | \theta_i) - \\ &Q(s, a | \theta_i) \nabla_{\theta_i} Q(s, a | \theta_i) - \\ &E_{s, a, r, s'}^N \rho(-1)^{1-B(a=y)} z(s, a) \nabla_{\theta_i} Q(s, a | \theta_i) - \\ &\rho E_{s, a, r, s'}^F \rho(-1)^{1-B(a=y)} z(s, a) \nabla_{\theta_i} Q(s, a | \theta_i) \end{aligned} \quad (20)$$

式 (20) 中, 第 3 项与少数类有关, 第 2 项与多数类有关。对于不平衡数据集 ($N > F$), 如果 $\rho = 1$, 则两个类别中样本的直接奖励值相同, 但多数类中的样本较多, 所以第 2 项中的数值要大于第 3 项, 训练出来的模型会偏向于多数类; 如果 $\rho = D_N/D_F$, 则 ρ 增加了少数类的立即回报, 加强了对网络损失函数的影响; 且本文所提模型充分考虑多数类和少数类的类间距和类内距, 使每个样本在保证不平衡比的前提下都有各自的奖励值。

2 滚动轴承振动数据不平衡的故障诊断及流程

滚动轴承振动数据不平衡的故障诊断及流程如图 3 所示。

深度强化学习中各个参数解释:

1) 状态 (State) S_i : 环境的状态取决于训练样本。在训练开始阶段, 智能体将第一个样本 x_1 作为初始状态 s_1 。在每个时间步长下, 状态 s_i 对应样本 x_i 。

2) 动作 (Action) A : 智能体的动作与训练集的标签相对应。在训练开始阶段, 智能体采取的行动是猜测类别标签。对于 10 分类问题, $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, 其中, $0 \sim 8$ 代表故障类即少数类, 9 代表正常类即多数类。

3) 回报 (Reward) R : 奖励是指来自环境的反馈, 通过该反馈可衡量智能体行为的优劣。为更好地指导智能体在不平衡数据集中学习最优的诊断策略, 整体上少数类样本的绝对奖励要高于多数类。当智能体正确或错误的识别少数类样本时, 环境反馈给智能体更大的奖励或

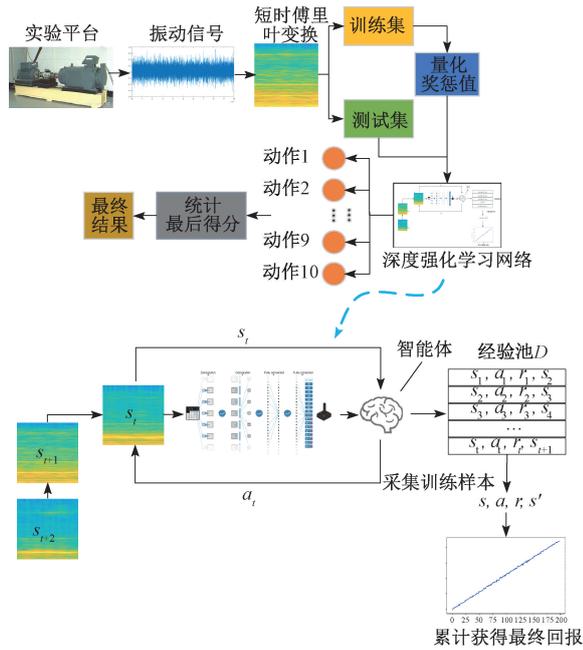


图 3 滚动轴承振动数据不平衡的诊断方法及流程
Fig. 3 The fault diagnosis method and process of imbalanced rolling bearing vibration data

惩罚。

4) 转移概率 (Transition probability) P : 该模型中, 转移概率 $P(s_{t+1} | s_t, a_t)$ 是确定的。智能体根据训练集中的样本顺序从当前状态 s_t 转移到下一状态 s_{t+1} 。

5) 探索率 (Discount factor) γ : $\gamma \in [0, 1]$ 平衡当前和未来的回报值。

6) 情节 (Episode): 强化学习中的情节指的是智能体根据某个策略从初始状态到最终状态这一过程。在该模型中, 设定从训练集样本中随机抽取一定数量, 智能体从第一个样本开始诊断到最后一个样本结束作为一个情节。

7) 策略 (Policy) π_θ : 策略 π_θ 指的是映射函数 $\pi: S \rightarrow A$, 其中, $\pi_\theta(s_t)$ 表示智能体在状态 s_t 下执行的动作 a_t 。在该模型中, 策略 π_θ 看作参数是 θ 的分类器。

数据不平衡下滚动轴承故障诊断流程:

1) 对经数据增强的训练集中滚动轴承多状态 (正常、内环不同故障程度、外环不同故障程度、滚动体不同故障程度共 10 类) 振动信号进行短时傅里叶变换 (short time Fourier transform, STFT) 变换, 获得二维时频域图像, 通过 K-means 方法确定各类样本的中心点, 利用欧式距离对每个样本赋予不同的回报值。

2) 为了对步骤 1) 中产生的数据进行一个甄别和筛选, 在通过 K-means 方法进行聚类的时候, 将聚类错误的样本剔除, 并再向模型申请一个新样本, 直到聚类正确为止。

3) 针对 DQN 网络中常见的稀疏奖励问题, 把每个时间步骤的经历 (s_t, a_t, r_t, s') 即当前轴承故障状态 s_t , 采取的故障诊断动作 a_t , 得到的立即回报 r_t 以及下一个状态, 当成记忆存储到经验池中, 这个记忆是一个固定长度的序列。在训练的时候, 从记忆池中随机采样, 计算更新模型。

4) 将滚动轴承的故障诊断过程模拟为 DQN 网络中的顺序决策过程, 当模型接收到振动信号的 STFT 图像, 智能体通过改进的残差网络提取深层特征, 然后依据改进的回报函数返回特定的奖励值。在迭代一定数目之后智能体将会如何诊断每张频谱图, 由于残差网络可以提取图片的深层特征, 所以对于测试集来说, 尽管像素分布有差别, 也会做出正确的识别动作。

5) 强化学习是不断试错的过程, 重要的是如何利用已学会的知识以及进行合理的试错, 这又称为探索-应用。深度强化学习继承了强化学习在未知环境下的适应力, 使得模型在多种不平衡比例下, 具有很好泛化性。在模型训练结束时, 为减少模型的探索随机性、使模型更加稳定, 借鉴迁移学习的思想, 将已训练的模型参数固定, 通过 Softmax 分类器进行最终的滚动轴承故障诊断。

训练的算法和环境模拟如算法 1 和 2 所示。

算法 1 环境模拟

Input: 训练数据 D 、奖励值 Z

Class *Enviroment*;

定义 *reset* 函数接口: 重置当前状态

def *reset*():

 return *step*(-1)

定义 *step* 接口: 返回下一张图片索引和奖励值

def *step*(*action*):

 return s', r

定义 *reward* 接口: 判断当前样本索引和 *action* 是否一致, 并返回奖励值

def *reward*(*action*)

 if *action* == y_t , then $r_t = z(s_t)$

 else $r_t = -z(s_t)$

算法 2 训练

Input: 训练数据 D , Episode K

初始化记忆回放单元 M

随机初始化参数 θ

初始化模拟环境 E

for epoch in epochs:

打乱训练集 D

for episode **in** K :

对每个状态用 ϵ -greedy 选择:

$$a_t = \text{spisode_greedy}(E, s_t)$$

$$s_{t+1}, r_t = E.\text{step}(a_t)$$

加入到经验池单元 M :

$$M.\text{append}(s_t, a_t, r_t, s_{t+1})$$

随机在经验池 M 内取本来进行训练

对损失函数 $L(\theta_i)$ 用梯度下降法进行更新:

$$L(\theta_i) = E_{s,a,r,s'}[(y_i - Q(s,a|\theta_i))^2]$$

end

end

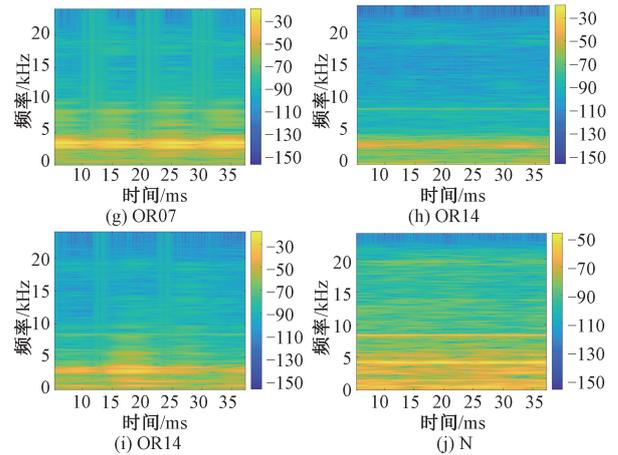


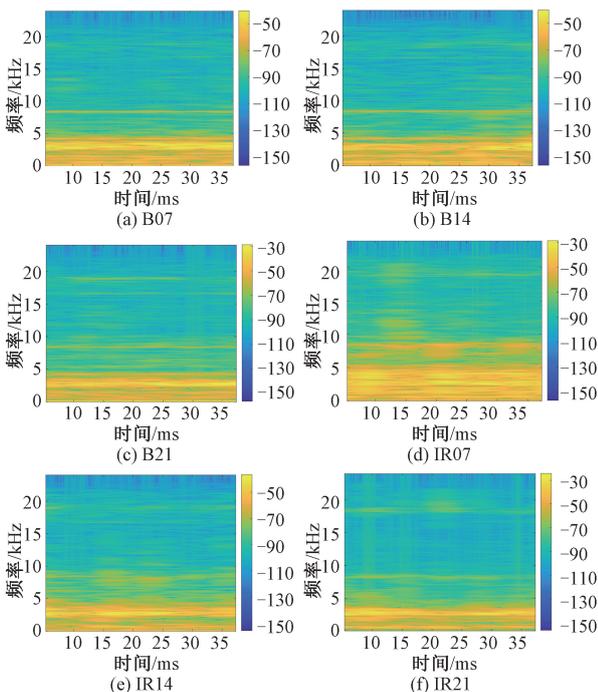
图4 不同状态下 STFT 结果

Fig. 4 STFT results in different states

3 实验分析

3.1 实验数据集设置

本文使用的滚动轴承振动数据由凯斯西储大学的轴承数据中心提供,这些数据由加速度传感器在 4 种负载条件下采集,采样频率为 12 和 48 kHz,负载类型包括 0、1、2 和 3 hp,根据负载的不同电机转速在 1 730 和 1 797 r/min 之间变化。振动信号包括 4 种不同的健康状态:正常条件(N)、外环故障(OR)、内环故障(IR)及滚动体故障(B),3 种缺陷故障直径分别为 0.007、0.014 和 0.021 inch。因此,该数据集包含 10 种轴承运行状态,以 0 hp 为例,每种状态下 STFT 结果如图 4 所示。



为了模拟真实情况中采集的振动数据分布不平衡,可能出现某一类样本数据多,另一类样本数据少。按照不同的不平衡比(imbalance ratio, IMR)来设置数据集进行多状态数据分布不平衡的实验。不平衡比指的是故障数据的数量与正常数据的数量之比。进行 10 种状态的不平衡和变负载实验,不平衡数据集设置如表 1 所示。

表1 不平衡实验数据集组成

Table 1 Composition of imbalanced experimental data set

IMR	1	2/3	1/2	1/10
B07	275	220	183	50
B14	275	220	183	50
B21	275	220	183	50
IR07	275	220	183	50
IR14	275	220	183	50
IR21	275	220	183	50
OR07	275	220	183	50
OR14	275	220	183	50
OR21	275	220	183	50
N	2 475	2 970	3 303	4 500

表 1 中不平衡比分别为 1、2/3、1/2、1/10,所有训练集样本数为 4 950 个。IMR = 1 时,滚动轴承正常状态与故障状态的样本数量均为总样本数量的 50%,此时数据集正常和故障的比例为 1:1,同时故障状态又分为 9 种;IMR = 2/3 时,滚动轴承故障状态的样本数量下降,正常状态的样本数量逐渐增多,此时训练集为轻度不平衡数据集;IMR = 1/2 时,正常状态的样本数量是故障状态的样本数量的 2 倍,此时训练集为中度不平衡数据集;

IMR = 1/10 时,正常状态的样本数量是故障状态的样本数量的 10 倍,此时训练集为极度不平衡数据集。为了便于比较,每种不平衡比下的测试集样本数目保持一致,为每种状态 50,共计 500 个测试样本。

3.2 模型结构及参数设置

模型用残差网络(Resnet-18)拟合 Q 函数,且对网络部分进行改进,基本参数设置如表 2 所示。

表 2 残差网络的参数设置

Table 2 Parameter setting of the residual network

编号	名称	核尺寸	核数量	输出尺寸	激活函数
1	输入层	-	-	227×227×3	-
2	卷积层 1	7×7	64	114×114×64	ReLU
3	最大池化层	-	-	57×57×64	ReLU
4	残差块 1	3×3	64	57×57×64	ReLU
5	残差块 2	3×3	64	57×57×64	ReLU
6	残差块 3	3×3	128	28×28×128	ReLU
7	残差块 4	3×3	128	28×28×128	ReLU
8	残差块 5	3×3	256	28×28×256	ReLU
9	残差块 6	3×3	256	28×28×256	ReLU
10	残差块 7	3×3	512	28×28×512	ReLU
11	残差块 8	3×3	512	28×28×512	ReLU
12	均值池化层	-	-	1×1×512	-
13	Flatten 层	-	-	512	-
14	全连接层	-	-	128	-
15	Q 值层	-	-	10	-
	动作输入层	-	-	10	-
16	Lambda 层	-	-	1	-

深度强化学习超参数设置:

智能体在与环境交互的过程中,探索概率 ε 从 1.0 据式(21)衰减至 $\varepsilon_{\min} = 0.01$ 。

$$\varepsilon = \max \left\{ \varepsilon_{\min}, 1 - \frac{(1 - \varepsilon_{\min}) \times \text{step}}{\text{total}} \right\} \quad (21)$$

式中:step 为当前迭代次数;total 为总迭代次数。探索率 $\gamma = 0.99$, 情节 $K = 512$, 迭代次数为 2 000, 回报值采用式(15)中的方法。

3.3 实验结果

为较全面的验证所提故障诊断模型在数据不平衡下的有效性及泛化能力,方便与其他文献的诊断结果进行对比,基于 Tensorflow 和 Keras 深度学习框架设置两种实验方案,GPU 用于加速计算。

1) 评价指标

在不平衡场景中,滚动轴承故障状态所占比例较少

(假设为 1%),如果模型将所有样本预测为正常状态,其准确率为 99%,但 G-mean 值为 0。可见准确率不能全面评价模型在不平衡数据下的表现;而 G-mean 值同时考虑样本的真阴率和真阳率,因此能衡量不平衡数据下模型的好坏,数值越大,模型的不平衡诊断性能越好。G-mean 的计算公式如式(22)所示。

$$G - \text{mean} = \sqrt{TP/(TP + FN)} \times \sqrt{TN/(TN + FP)} \quad (22)$$

式中:TP 为真正例;FN 为假反例;TN 为真反例;FP 为假正例。

针对 G-mean 指标仅能评价不平衡二分类的问题,文献[19]提出一种不平衡的多分类指标 G-meantotal,即从众多类别中任取两个类别 C_i 和 C_j ,计算两类样本分类结果的 G-mean(C_i, C_j) 指标,再对所有的 G-mean 值加权求和,如式(23)所示。

$$G - \text{mean}_{\text{total}} = \frac{2}{K \times (K - 1)} \times \sum_{1 \leq i < j \leq K} G - \text{mean}(C_i, C_j) \quad (23)$$

本文实验涉及平衡和非平衡数据集,所以采用准确率和 G-meantotal 作为评价指标。

2) 实验方案 1

为验证所提模型在数据不平衡下的诊断能力,在滚动轴承工作在相同负载的情况下,对 4 种不同不平衡比例的训练集进行训练。测试集如 3.1 节所述,总计 500 个测试样本;与其他 3 种方法进行对比的结果如图 5 所示,其中,横坐标的 0 负载指的是训练和测试都是采用的同种负载(0 负载),其他 123 负载类同;SMOTE 方法为传统过采样方法。

根据图 5 的准确率对比可看出经过改进的 DQN 在数据集不平衡下表现出了较大的优势。IMR = 1 时,改进后的 DQN 模型较改进前的准确率提高了 1%,较传统的 SMOTE 方法提高了 5%~8%;当 IMR = 2/3 不平衡比增加时,改进后的模型诊断结果整体高于 99%,甚至达到 100%,较改进前提高了 1%左右,较传统方法提高了 6%~8%;即使当 IMR = 1/10 的极度不平衡时,故障诊断准确率也集中在 99%~100%之间,优于 SMOTE 和 CNN 方法。而 CNN 方法在不平衡测试中,只能诊断出训练集中的多数类数据,在数据集不平衡下表现不佳。

因为数据是不平衡的,单独使用分类准确率来评价模型的性能不够全面,现对比实验结果的 G-meantotal 指标。同时为了对比更加明显,现仅对那些改进后模型准确率低于 99%的数据类别进行 G-meantotal 指标的对比。如图 6 所示,0 负载、IMR = 1 时,G-meantotal 指标高于改进前的模型和 SMOTE、CNN 方法,进一步说明本文模型具有优秀的不平衡数据的诊断能力。

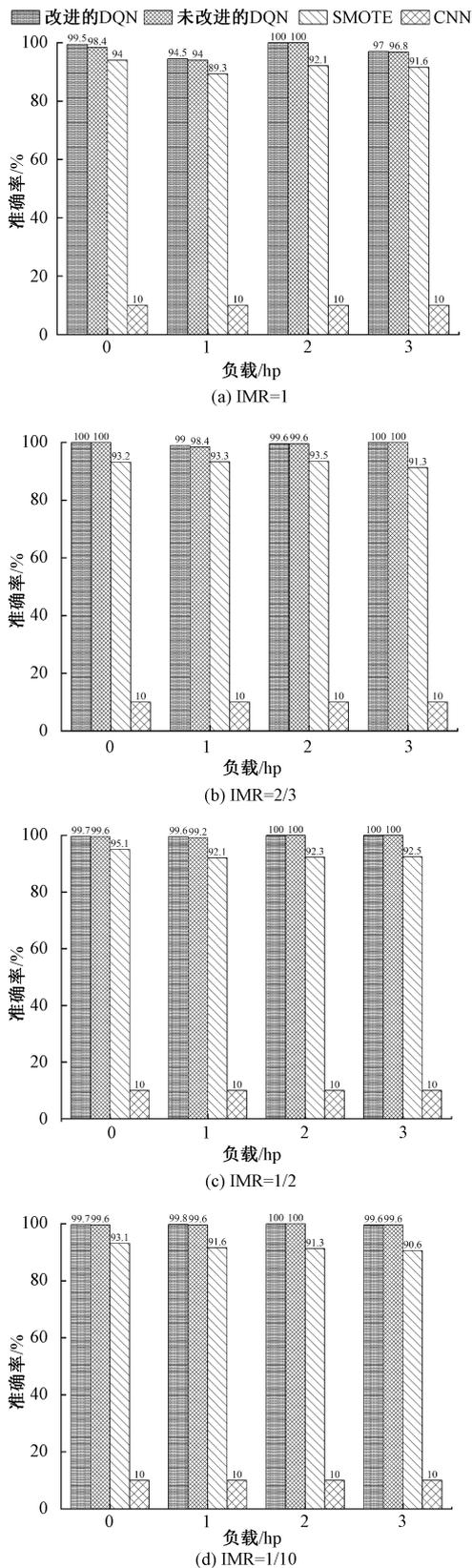


图5 不同平衡比的4种方法准确率比较(实验方案1)
Fig. 5 Comparison of accuracy of four methods with different balance ratios (experiment 1)

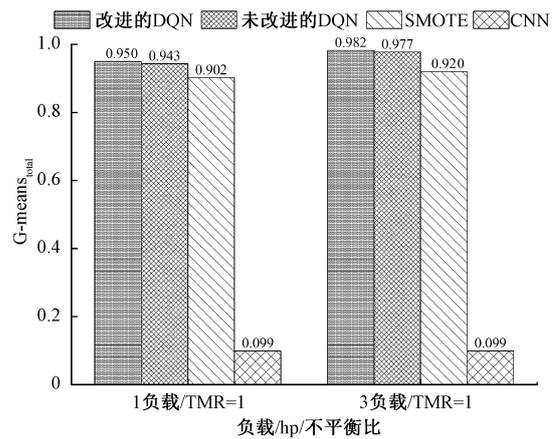


图6 G-means_{total} 指标部分结果(实验方案1)
Fig. 6 Results of some G-means_{total} indicators (experiment 1)

图5 准确率和图6 G-means_{total} 从实验结果的角度共同验证了模型在不平衡数据下的有效性。对不平衡比例为1,0负载的训练回报值曲线进行分析,根据1.2节中回报值的设定,模型每迭代一次均以经验池中回放512个样本,且都符合不平衡比例为1的分布。当训练结束时,本文改进的DQN模型如果处于最佳状态,回报值应该集中在 $512/2 \times 9 + 512/2 \approx 2560$ 左右,如图7所示。

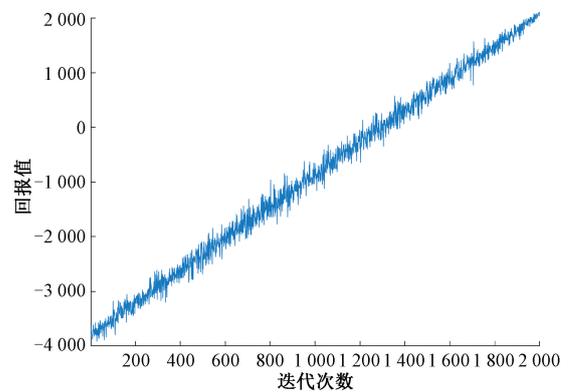


图7 训练回报值
Fig. 7 Training reward value

为更直观地证明所提方法的有效性,以0负载、IMR=1为例,采用t-分布邻域嵌入(t-distribution stochastic neighbor embedding, t-SNE)算法^[20],将改进前后的DQN网络最后一层全连接层的特征降维至二维并以散点图形式表示出来,如图8和9所示。

通过对图8观察可得,改进前的模型存在错分及某些类别决策边界不明显的问题(矩形标注),具有一定的提升空间;经过改进后的模型,如图9所示,决策边界明显,错分现象较少,提高了故障诊断的准确率。

3) 试验方案2

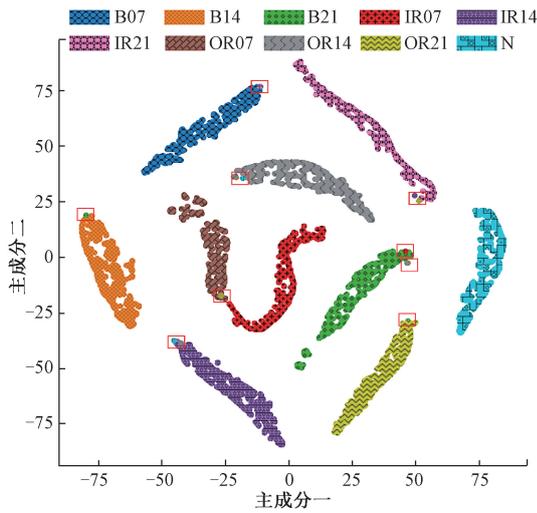


图 8 改进前 DQN 网络特征可视化图

Fig. 8 Diagram of the unimproved DQN network feature visualization

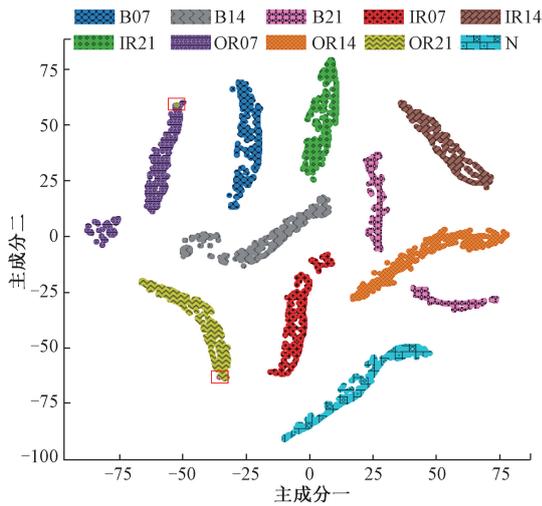
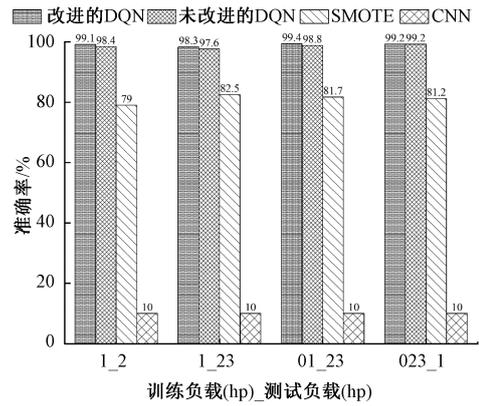


图 9 改进后 DQN 网络特征可视化图

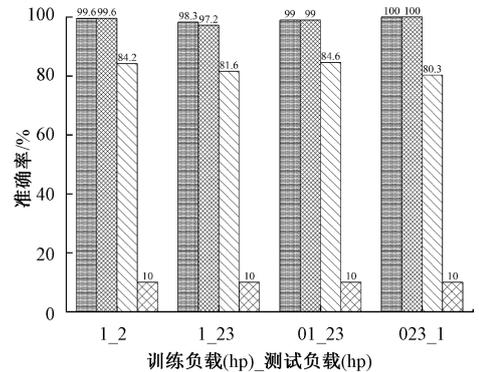
Fig. 9 Diagram of the improved DQN network feature visualization

实验方案 1 的实验结果展现了本文模型可以很好地解决振动数据分布不平衡问题,为验证本文所提模型的泛化能力,对变负载且不平衡下的数据对模型进行测试,实验结果如图 10 所示。

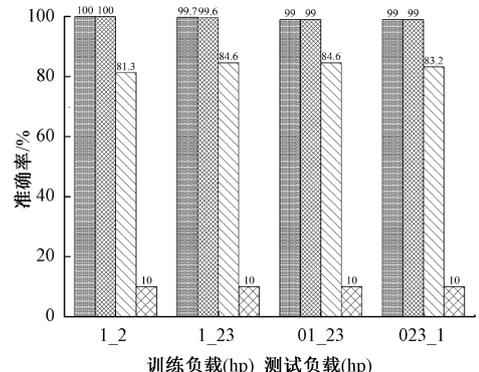
图 10 为 4 种不平衡比的滚动轴承变负载实验分类准确率,可以看出,无论训练集和测试集是单一负载还是多种负载,本文方法均具有良好的分类准确率。以 01_23 数据进行实验,当 $IMR=1$ 时,分类准确率为 99.4%,当 $IMR=1/2$ 时,分类准确率也能有 99% 的准确率,远高于 SMOTE 方法;以 023_1 数据进行实验,当 $IMR=1$ 时,分类准确率达 99.2%。即使在 $IMR=1/10$



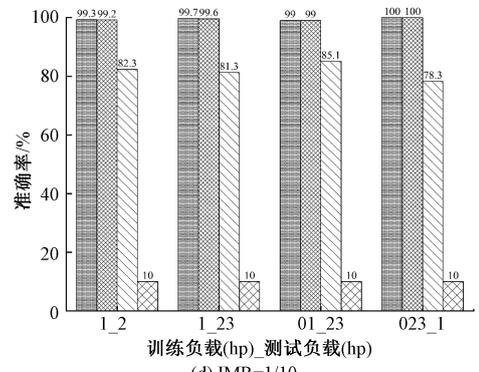
(a) $IMR=1$



(b) $IMR=2/3$



(c) $IMR=1/2$



(d) $IMR=1/10$

图 10 不同平衡比 4 种方法准确率比较 (实验方案 2)

Fig. 10 Comparison of accuracy of four methods with different balance ratios (experiment 2)

的极度不平衡情况下,分类准确率能达到 100%,比 SMOTE 方法更加优秀。

图 11 所示为不平衡指标 $G\text{-means}_{\text{total}}$ 下,以 01_23 数据为例,当 $\text{IMR} = 1$ 和 $\text{IMR} = 2/3$ 时, $G\text{-means}_{\text{total}}$ 值达到 0.988 和 0.989 优于传统方法和改进前的方法。

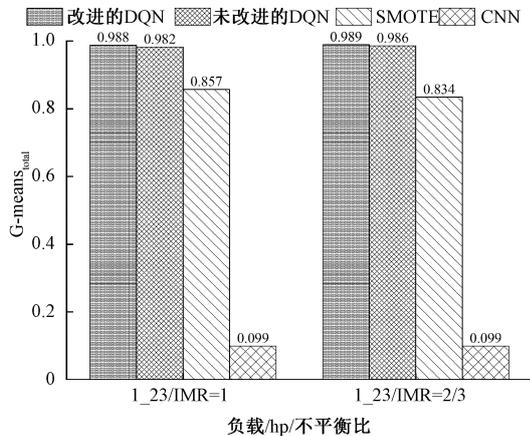


图 11 $G\text{-means}_{\text{total}}$ 指标部分结果(实验方案 2)

Fig. 11 Results of some $G\text{-means}_{\text{total}}$ indicators (experiment 2)

为直观地证明所提方法的有效性,现以 023 负载训练 1 负载测试, $\text{IMR} = 1/10$ 为例,用 t-SNE 方法展示本文模型在不平衡且变负载下的故障诊断可视化,如图 12 所示。

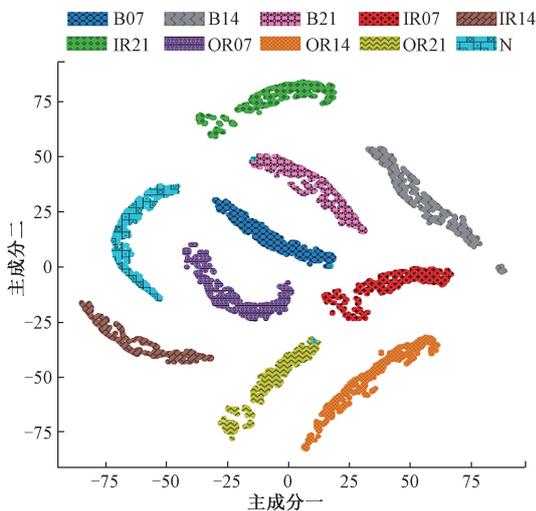


图 12 改进后 DQN 网络特征可视化图

Fig. 12 Diagram of improved DQN network feature visualization

从图 12 中可以看出,在不平衡且变负载下仅有极个别的样本未能正确检测,具有优秀的泛化能力。

可见,即使在数据极度不平衡的情况下,本文所提方

法能很好地适应多变的环境,即使数据不平衡、负载发生变化,也较未改进前具有更高的准确率。因此,本文所提出的方法可以很好地解决振动数据正常状态与故障状态的数据分布不平衡且变负载问题。

4 结 论

1) 提出一种结合 K-means 和空间距离构建强化学习中回报函数的方法,对每个样本的回报值进行量化,获得更为准确的回报值;提出通过修改后的 Resnet-18 网络优化原有的 DQN 模型,使原有模型更深,泛化性更好,提高了模型稳定性和诊断准确率。

2) 改进的 DQN 模型可很好地解决振动数据正常状态与故障状态的数据分布不平衡问题,不平衡且变负载情况下的 $G\text{-meantotal}$ 得分达到了 0.982 左右,表现优异。当数据极度不平衡时准确率达到 97%~99%,部分达到 100%,与传统的平衡分类方法对比提高了 5%~8%。

3) 综上实验结果,所提模型在不平衡和变负载下表现稳定,整体性能优于改进前的 DQN 网络和传统的 CNN 网络,且优于原有的基于数据的不平衡解决办法,呈现了深度强化学习的另一种思维方式,以及深度强化学习作为一种成熟学习框架的潜力。

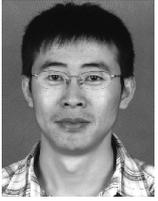
下一步工作将进一步研究在数据不平衡下的变工况滚动轴承故障诊断及深度强化学习网络中参数优化问题。

参考文献

- [1] 张淑清,苑世钰,姚玉永,等. 基于 ICEEMD 及 AWOA 优化 ELM 的机械故障诊断方法[J]. 仪器仪表学报, 2019, 40(11): 172-180.
ZHANG SH Q, WAN SH Y, YAO Y Y, et al. Machinery fault diagnosis method based on ICEEMD and AWOA optimized ELM [J]. Chinese Journal of Scientific Instrument, 2019, 40(11): 172-180.
- [2] 包萍,刘运节. 不平衡数据集下基于生成对抗网络的改进深度模型故障识别研究[J]. 电子测量与仪器学报, 2019, 33(3): 176-183.
BAO P, LIU Y J. Research on fault identification based on improved deep model in combination of generative adversarial networks under unbalanced data sets [J]. Journal of Electronic Measurement and Instrument, 2019, 33(3): 176-183.
- [3] 宫文峰,陈辉,张美玲,等. 基于深度学习的电机轴承微小故障智能诊断方法[J]. 仪器仪表学报, 2020, 41(1): 195-205.

- GONG W F, CHEN H, ZHANG M L, et al. Intelligent diagnosis method for incipient fault of motor bearing based on deep learning [J]. Chinese Journal of Scientific Instrument, 2020, 41(1): 195-205.
- [4] 董勋, 郭亮, 高宏力, 等. 代价敏感卷积神经网络: 一种机械故障数据不平衡分类方法[J]. 仪器仪表学报, 2019, 40(12): 205-213.
- DONG X, GUO L, GAO H L, et al. Cost sensitive convolutional neural network: A classification method for imbalanced data of mechanical fault[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 205-213.
- [5] 黄海松, 魏建安, 任竹鹏, 等. 基于失衡样本特性过采样算法与SVM的滚动轴承故障诊断[J]. 振动与冲击, 2020, 39(10): 65-74, 132.
- HUANG H S, WEI J AN, REN ZH P, et al. Rolling bearing fault diagnosis based on imbalanced sample characteristics oversampling algorithm and SVM [J]. Journal of Vibration and Shock, 2020, 39(10): 65-74, 132.
- [6] WANG H, KE Y, LUO G, et al. Compressed sensing of roller bearing fault based on multiple down-sampling strategy [J]. Measurement Science and Technology, 2015, 27(2): 025009.
- [7] MAO W, HE L, YAN Y, et al. Online sequential prediction of bearings imbalanced fault diagnosis by extreme learning machine[J]. Mechanical Systems and Signal Processing, 2017(83): 450-473.
- [8] JIA F, LEI Y, LU N, et al. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization [J]. Mechanical Systems and Signal Processing, 2018, 110: 349-367.
- [9] 陈仁祥, 吴昊年, 杨黎霞, 等. 多分类器集成加权均衡分布适配的滚动轴承寿命阶段识别[J]. 仪器仪表学报, 2019, 40(10): 66-73.
- CHEN R X, WU H N, YANG L X, et al. Rolling bearing life stage recognition based on multi-classifier integration of the weighted and balanced distribution adaptation[J]. Chinese Journal of Scientific Instrument, 2019, 40(10): 66-73.
- [10] WU Z, GUO Y, LIN W, et al. A weighted deep representation learning model for imbalanced fault diagnosis in cyber-physical systems[J]. Sensors, 2018, 18(4): 1096.
- [11] LIN E, CHEN Q, QI X. Deep reinforcement learning for imbalanced classification [J]. Applied Intelligence, 2020: 1-15.
- [12] JANISCH J, PENVNY T, LISY V. Classification with costly features using deep reinforcement learning [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 3959-3966.
- [13] FENG J, HUANG M, ZHAO L, et al. Reinforcement learning for relation classification from noisy data [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 5779-5786.
- [14] WIERING M A, VAN HASSELT H, PIETERSMA A D, et al. Reinforcement learning algorithms for solving classification problems [C]. 2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), IEEE, 2011: 91-96.
- [15] DING Y, MA L, MA J, et al. Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach [J]. Advanced Engineering Informatics, 2019 (42): 100977.
- [16] FAN J, WANG Z, XIE Y, et al. A theoretical analysis of deep Q-learning. learning for dynamics and control [C]. PMLR, 2020: 486-489.
- [17] WU S, ZHONG S, LIU Y. Deep residual learning for image steganalysis [J]. Multimedia Tools and Applications, 2018, 77(9): 10437-10453.
- [18] 张龙, 宋成洋, 邹友军, 等. 基于VMD多特征融合与PSO-SVM的滚动轴承故障诊断[J]. 机械设计与研究, 2019, 35(6): 96-104.
- ZHANG L, SONG CH Y, ZOU Y J, et al. A rolling bearing fault diagnosis method based on VMD multi feature fusion and PSO-SVM [J]. Machine Design & Research, 2019, 35(6): 96-104.
- [19] ESPINDOLA R P, EBECKEN N F F. On extending f-measure and g-mean metrics to multi-class problems [J]. WIT Transactions on Information and Communication Technologies, 2005(35): 25-34.
- [20] ZHENG J, JIANG Z, PAN H. Sigmoid-based refined composite multiscale fuzzy entropy and t-SNE based fault diagnosis approach for rolling bearing [J]. Measurement, 2018(129): 332-342.

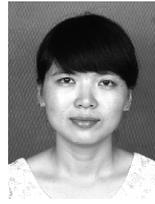
作者简介



康守强, 2011 年于白俄罗斯国立大学获得博士学位, 现为哈尔滨理工大学教授、博士生导师, 主要研究方向为非平稳信号处理, 机械故障诊断、状态评估与预测技术。

E-mail: kangshouqiang@163.com

Kang Shouqiang received his Ph. D. degree from Belarusian State University in 2011. He is currently a professor and a Ph. D. advisor at Harbin University of Science and Technology. His main research interests include non-stationary signal processing, machinery fault diagnosis, condition assessment and prediction technology.



王玉静 (通信作者), 2016 年于哈尔滨工业大学获得博士学位, 现为哈尔滨理工大学副教授、硕士生导师, 主要研究方向为非平稳信号处理, 机械故障诊断、状态评估与预测技术。

E-mail: mirrorwyj@163.com

Wang Yujing (Corresponding author) received her Ph. D. degree from Harbin Institute of Technology in 2016. She is currently an associate professor and a master advisor at Harbin University of Science and Technology. Her main research interests include non-stationary signal processing, machinery fault diagnosis, condition estimation and prediction technology.